Cadernos de
ESTUDOS LINGÜÍSTICOS – (56.1), Campinas, Jan./Jun. 2014

PHONOLOGICAL REGULARITIES AMONG DEFECTIVE VERBS

ANDREW NEVINS
UCL/UFRJ
GEAN DAMULAKIS
UFRJ
MARIA LUISA FREITAS
UNICAMP

RESUMO: Este artigo discute regularidades fonológicas dos verbos defectivos em português, a partir de um estudo experimental realizado com intuito de averiguar duas questões: primeiro, se os verbos descritos como defectivos (Cunha & Cintra 2013) são concebidos como tal por falantes nativos; e segundo; se a generalização de Postma (2013) de que a grande maioria dos verbos defectivos em português apresenta uma sonorante coronal (n, l, r, ɲ) seguindo a última vogal da raiz verbal é ou não internalizada pelos falantes.
Palavras chave: verbos defectivos, generalização de Postma, design experimental

RESUMEN: Este artículo discute regularidades fonológicas de los verbos defectivos en portugués, a partir de un estudio experimental llevado a cabo con intento de investigar dos cuestiones: primero, si los verbos descritos como defectivos (Cunha & Cintra 2013) son concebidos como tal por hablantes nativos; segundo; si la generalización de Postma (2013) de que la gran mayoría de verbos defectivos en portugués tiene una sonorante coronal (n, l, r, ɲ) después de la última vocal de la raíz verbal es o no internalizada tiene por los hablantes.
Palabras clave: verbos defectivos, generalización de Postma, design experimental

1. INTRODUCTION

Defectiveness can be defined as when the morphosyntactic space is not fully realized by the exponent space. Spanish and Portuguese are well-known for their 'defective' verbs: those which possess infinitive, participle, and preterite forms, but a limited set of present tense forms. One such example is Portuguese *falir* 'to fail', which possesses the infinitive, the participles *falido* and *falindo*, the full set of preterit and future forms, but no present tense indicative except for the 1pl *falimos,* and no present tense subjunctive forms whatsoever. A notable generalization about the 'missing' forms is that they are in a large part

'rhizotonic', i.e. they would have stress on the root, rather than on an inflectional desinence (similar, perhaps, to defective nouns in Russian). Nonetheless, this generalization is not the whole story, as the would-be 1pl subjunctive form *falamos* would bear stress on the penultimate syllable (and not the root), and thereby the correct generalization seems to be that the only 'effable' forms are those which overtly retain the theme vowel /i/ after the root. The preceding statement thus implicitly carries one of the generalizations at hand: defective verbs in Portuguese are all in the 3rd conjugation.

Characterizing the set of verbs which show this defectivity (as clearly, not all 3rd conjugation verbs are defective) has proved an elusive task for many grammarians. While some of them might be classified in terms of semantic restrictions (e.g. the putatively defective verb *latir* 'to bark' would sound odd or uncivilized in the 1sg), the fact that such forms are also defective in say, the 3rd person subjunctive rules out any appeal to semantics alone. In characterizing the set of defective verbs, an important caveat is also to make sure that speakers *know* the verbs at all; those which are so low-familiarity as to yield uncertainty about any of the forms should be excluded from the first pass of arriving at a generalization. One of the often tempting explanations is based on homophony avoidance, as the three defective verbs below would have a 1sg indicative and subjunctive identical to that of another, non-3rd conjugation, and high-familiarity verb:

(1)

a. *polir*[1] 'polish': 1sg.ind *pulo,* 1sg.sbj *pula* (forms confusable with *pular* 'jump')

b. *parir* 'give birth': 1sg.ind *paro,* 1sg.sbj *para* (forms confusable with *parar* 'stop')

c. *falir* 'fail': 1sg.ind *falo,* 1sg.sbj *fala* (forms confusable with *falar* 'talk')

This, too, cannot be the whole story, as the set of 40 defective verbs compiled in the Aurélio dictionary and Cunha & Cintra's grammars contain many which do not submit to this explanation. Perhaps, indeed, we are relegated to a position no more intriguing than that of Maiden & O'Neill (2010), who state 'The major domain of defectiveness (present subjunctive and 1sg present indicative) seems, and is, irreducibly arbitrary if one seeks a motivation outside the morphological system itself.'

One general intuition shared by many works (Maiden & O'Neill 2010, among others) is that defective forms are those which not only lack specific allomorphs, but which seem to actively *resist* certain kinds of otherwise-available allomorphy.

[1] Traditionally, the verbs 'polir' and 'parir' have been described as regular verbs. However, as we will show in our results, there is strong evidence that these verbs may be considered truly defectives.

Focusing on the Iberian defective verbs, Maiden & O'Neill (2010:109) point out that 'defective verbs of learnèd origin never show any kind of allomorphy' – not even where it might be expected, e.g. as in the preterite forms of Spanish *abolió* (not **abulió).* They note that many of the defective verbs in Spanish and Portuguese are not attested before the 16th century, and that features such as retention of the intervocalic /l/ in defective *demolir* 'to demolish' attest to this fact. This claim is empirically testable, and likely to be disconfirmed if wug tests show that speakers treat certain verbs as defective and others as non-defective even without indication of their neo-classical origins.

How do speakers know that they cannot produce certain forms, if the forms they have never produced have never been indicated as ungrammatical during acquisition? This would seem to instantiate a core case of the poverty of the stimulus problem in morphology, and like all such problems, a principled generalization might be able to guide learners in delimiting the space of ungrammatical structures without having explicit negative evidence for their status.

In fact, a phonological generalization has recently been proposed, at least for Portuguese. Postma (2013) noticed that of the 40 defective verbs listed in Aurélio dictionary and Cunha & Cintra, 35 (87%) have a coronal sonorant (n,l,r,ɲ) immediately following the stem-final vowel (e.g. *colorir, banir*, *polir*). In terms of the explanation Postma offers, his idea is that deleted theme-vowel harmony (e.g. *dormir, dormes, durmo, durmas*; Harris 1974) competes with the coronal for positioning within the root. While we must refer the reader to Postma's work for the details of the proposal, the intuition is that the fact that defective verbs are defective precisely where the theme vowel /i/ cannot be overtly realized has an interaction with the presence of a coronal sonorant following the final root-vowel.

The question we therefore wish to examine is whether indeed, speakers internalize such a generalization. In fact, before doing wug-experiments to test the productivity of this putative generalization, we need to know what the actual list of defective verbs *is* for real speakers. For example, are *explodir* 'explode' and *latir* 'bark' really defective (and hence exceptions to Postma's generalization)? In order to conduct this study, we took all 40 defective verbs and removed those that two graduate students did not know the meanings of. This left 21 verbs. We asked participants to rate familiarity of each verb from 1-5 in a pretest and excluded those with a mean lower than 2. Participants (N=30) produced the finite 1pl, 3sg indicative and the 3sg subjunctive, and then rated their confidence in their own productions, following the procedure of Albright (2003, 2009), who used the measure of a speaker's confidence in his/her own production as a measure of defectiveness.

The rest of this paper is organized as follows. In Section 2, we discuss the experimental design; in Section 3, we present the main results of the statistical analyses; and in Section 4, we conclude the article.

2. EXPERIMENTAL DESIGN

Our experiment was composed of three parts, and represents an adapted version of the experimental procedure presented in Albright (2003, 2009). In the first part, the speakers had to rate the familiarity of 21 verbs listed as defective in Cunha & Cintra (2013) and Aurélio Dictionary from a *Likert scale* of 5 points, indicating the level of their knowledge in relation to each verb, in the following manner:

(2) Likert scale for Task 1:

1 – I do not know this verb; I have never used or heard it.

2 – I do not know this verb well; I have never used it, but I have heard it before.

3 – I know this verb, I use it and I have heard it sometimes.

4 – I know this verb well; I use it and have heard it frequently.

5 – I know this verb very well; I use it and have heard it a lot.

In addition to the list of 21 verbs, six verbs were used as training items in preparation for the test. Tables 1 and 2 below show the verbs tested, where Table 2 includes 'wug' verbs (e.g. non-existent, invented verbs) as part of the training sequence in use of the scale.

Table 1: Tested verbs

abolir	'to abolish'	engolir	'to swallow'	imergir	'to drown'
balir	'to bleat'	exaurir	'to exhaust'	jungir	'to merge'
banir	'to ban'	explodir	'to explode'	latir	'to bark'
brandir	'to brandish'	falir	'to fail'	parir	'to give birth'
colorir	'to tint'	florir	'to flourish'	polir	'to polish'
demolir	'to demolish'	ganir	'to yelp'	ruir	'to collapse'
emergir	'to emerge'	grunhir	'to grunt'	ungir	'to anoint'

Table 2: Training items

pantominar	wug-word
sabujar	wug-word
tabizar	wug-word
tirar	'to take out'
vapular	wug-word
velejar	'to sail'

The second part of the experiment consisted of a fill-in-the-blank production task (Albright, 2003:6), in which the speakers were asked to produce three inflected forms of each verb from its infinitive: 3rd person singular (3sg.ind), 1st person plural indicative (1pl.ind) and 3rd person singular subjunctive (3sg.subj). The 1pl.ind form is non-rhizotonic and hence expected to be grammatical just like the infinitive, whereas the 3sg.ind and 3sg.subj forms are rhizotonic and hence expected to have low ratings if the verbs are defective. Importantly, our study (unlike others, including Albright's) specifically focused on the defectiveness of 3sg subjunctive, in order to see how truly general the pattern was, and avoid the potential confound of 1sg semantic restrictions.

(3) Task 2: frame example

O cachorro do Mateus adora latir. Ele _____ muito durante a noite. Espero que hoje ele não _____ muito. Quando passamos pelo portão, nós _____ para ele.

Mateus' dog loves to bark. He _____ a lot during the night. I hope he does not _____ a lot today. When we pass by the gate, we _____ at him.

In the third part of the experiment, each participant had to rate the forms that he or she produced, on a *Likert scale* from 1 to 5 rating the preference on their own production, as we show below:

(4) Likert scale for Task 3:

1. I really dislike this form.

2. I dislike this form.

3. I do not have a strong preference.

4. I like this form.

5. I really like this form.

We applied six different versions of the experiment in order to randomly distribute the three person/number/mood inflections tested in all six possible orders along different frames for Task 2. A total of 30 participants took part in the experiment, all of whom were native speakers of Brazilian Portuguese from Rio de Janeiro with college-level education. In the next section, we discuss the data treatment and the main results of the statistical analyses.

3. RESULTS AND ANALYSES

From the results of the experiment, we created a spreadsheet using the software *SPSS Statistic Data Editor*. We considered the results of Tasks 1 and 3 on a scale of 1 to 5 and calculated the means of each verb in both tasks, as shown in Tables 3 and 4 below.

Table 3: Descriptive statistics for Task 1

Verb		Mean	Std. Deviation
abolir	'to abolish'	3.60	.72
balir	'to bleat'	1.47	.62
banir	'to ban'	3.70	.91
brandir	'to brandish'	2.00	.97
colorir	'to tint'	4.50	.72
demolir	'to demolish	4.47	.77
emergir	'to emerge'	3.60	.99
engolir	'to swallow'	4.80	.48
exaurir	'to exhaust'	2.47	1.09
explodir	'to explode'	4.80	.54
falir	'to fail'	4.33	.79
florir	'to flourish'	3.40	1.12
ganir	'to yelp'	1.70	.59
grunhir	'to grunt'	2.57	.77
imergir	'to drown'	3.00	1.04
jungir	'to merge'	1.13	.34
latir	'to bark'	4.40	.96
parir	'to give birth'	3.93	.97
polir	'to polish'	3.40	1.09
ruir	'to collapse'	3.23	.89
ungir	'to anoint'	3.57	1.24

From the data above, we may conclude that in the familiarity rating task (Task 1), the verbs *balir* 'to bleat', *ganir* 'to yelp', *jungir* 'to merge' and *brandir* 'to brandish' are the least familiar verbs, exhibiting means equal or below a score of 2. We now turn to the ratings in Task 3.

Table 4: Descriptive statistics for Task 3

Verb		3Sbj		1Pl		3Sg	
		Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
abolir	'to abolish'	2.20	1.22	3.57	1.17	3.00	1.26
balir	'to bleat'	1.21	.50	1.61	.83	1.38	.78
banir	'to ban'	1.89	1.15	3.52	1.01	2.95	1.21
brandir	'to brandish'	1.93	1.09	1.93	.83	1.83	.98
colorir	'to tint'	2.28	1.44	4.17	1.09	3.71	1.16
demolir	'to demolish'	2.04	1.22	3.71	1.45	3.20	1.55
emergir	'to emerge'	2.61	1.17	3.00	1.35	3.29	1.16
engolir	'to swallow'	4.03	.93	4.43	.69	4.54	.66
exaurir	'to exhaust'	2.17	1.05	2.37	1.00	2.36	1.04
explodir	'to explode'	4.40	.89	4.45	.67	4.54	.71
falir	'to fail'	2.08	1.20	3.67	1.32	3.43	1.51
florir	'to flourish'	2.36	1.33	3.09	1.31	2.26	1.37
ganir	'to yelp'	1.50	.78	1.67	.84	1.50	.75
grunhir	'to grunt'	1.73	1.05	2.10	1.11	1.93	1.10
imergir	'to drown'	1.96	1.23	2.86	1.24	3.07	1.28
jungir	'to merge'	1.03	.19	1.15	.37	1.12	.33
latir	'to bark'	3.27	1.41	3.76	1.24	4.47	1.01
parir	'to give birth'	1.79	.94	3.45	1.24	2.77	1.36
polir	'to polish'	1.62	1.08	3.25	1.35	2.10	1.17
ruir	'to collapse'	1.48	.69	2.69	1.32	2.18	1.33
ungir	'to anoint'	2.59	1.55	3.21	1.32	2.84	1.31

Focusing specifically on the 3sg.subj., we performed a Spearman correlation test, in order to verify if there is a correlation between the results for Task 1 and Task 3. As we can notice in Table 5 below, there is a positive correlation between the variables, meaning that the higher the score in Task 1, higher the score in Task 3, i.e., if the verb is more familiar, it will increase the tendency of its inflected form to be well rated.

Table 5: Spearman Correlation - Test 1 and 3sg.subj Test 3

Spearman's rho	Correlation Coefficient	**.493**
	Sig. (2-tailed)	**.000**
	N	**592**

Turning to the central question at hand, we conducted an Independent Samples t-test in order to determine if there is a difference between the means of Task 3 results for the 3sg.subj and 1pl.ind and, consequently, to infer the defectiveness of each verb for the speakers. The null hypothesis (H_0) is that there is no statistically significant difference between the mean of the two groups. To perform the t-test, the SPSS software includes the Levene's test of Equality of Variances. This test shows if there is homogeneity variance between the tested groups. The results for both tests are shown below. We will present the shown least familiar verbs separately, and will not consider these verbs in our analysis.

Table 6: Levene's Test and Independent Samples t-test

	Verb		Levene's Test for Equality of Variances		t-test for Equality of Means				
			F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. Error Difference
Equal variances assumed	abolir	'to abolish'	.003	.959	-4.232	53	.000	-1.367	.323
	banir	'to ban'	.006	.937	-5.510	52	.000	-1.630	.296
	colorir	'to tint'	2.031	.160	-5.304	51	.000	-1.891	.357
	demolir	'to demolish	2.283	.139	-4.077	41	.000	-1.667	.409
	emergir	'to emerge'	1.598	.212	-1.126	50	.266	-.393	.349
	engolir	'to swallow'	.618	.435	-1.830	56	.073	-.395	.216
	exaurir	'to exhaust'	.054	.817	-.754	58	.454	-.200	.265
	explodir	'to explode'	1.064	.307	-.240	50	.811	-.055	.227
	falir	'to fail'	.664	.420	4.329	45	.000	-1.590	.367
	florir	'to flourish'	.176	.677	-1.831	42	.074	-.727	.397
	grunhir	'to grunt'	.008	.927	-1.315	57	.194	-.370	.281
	imergir	'to drown'	.072	.790	-2.705	54	.009	-.893	.330
	latir	'to bark'	.662	.419	-1.418	57	.162	-.492	.347
	parir	'to give birth'	3.588	. 063	-5.723	56	.000	-1.655	.289
	polir	'to polish'	1.863	.178	5.034	55	.000	-1.629	.324
	ungir	'to anoint'	1.345	.251	-1.648	55	.105	-.628	.381
Equal variances not assumed	ruir	'to collapse'	10.225	.002	-4.191	36.7	.000	-1.210	.289

Table 7: Levene's Test and Independent Samples t-test for the less familiar verbs

	Verb		Levene's Test for Equality of Variances		t-test for Equality of Means				
			F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. Error Difference
Equal variances assumed	ganir	'to yelp'	.480	.491	-.796	58	.429	-.167	.209
Equal variances not assumed	balir	'to bleat'	6.685	.012	-2.144	44.2	.038	-.393	.183
	brandir	'to brandish'	4.621	.036	.010	50.4	.992	.003	.260
	jungir	'to merge'	11.093	.002	-1.492	36.0	.144	-.119	.080

There are four verbs (*balir, brandir, jungir, ruir*) that have violated the assumption of homogeneity of variances (p-value lower than 0.05 for the Levene's test), as we can see at the bottom of both tables. In such cases, we used pooled variance in order to perform the t-test.

As the results show, forms in red are significantly different in their ratings of the non-rhizotonic 1pl.ind form and the rhizotonic 3sg.subj form, allowing us to reject the null hypothesis. The verbs *abolir, banir, colorir, demolir, falir, imergir, parir, polir and ruir* are in fact defectives. Almost all forms that were defective have a coronal sonorant in the root, upholding Postma's generalization. The only two exceptions are the verbs *imergir* 'to drown' and *ruir* 'to collapse'.

Importantly, the verbs that are defective according to these results can have a mid vowel, a high vowel, or a low vowel in the root, thereby rendering the generalization about the nature of the consonant more predictive than about, say, having mid-vowels (an explanation often proposed in the literature). In addition, Postma points out that defective verbs usually have the structure: vR(C)ir, where v is any vowel (not only mid-vowels); R is any coronal sonorant and (C) is an optional consonant. Our results show, however, that this optional consonant is not salient/productive in delimiting the class of defective verbs. The verbs that are robustly defective are those that have the structure vRir.

These results are in accord with the findings of Damulakis & Rodrigues (2013), who found that in medieval Portuguese, verbs such as *ferir* (with modern 1sg *eu firo*) had occasional alternative 1sg forms such as *feyro*, suggesting a metathetic migration of the theme vowel to the tonic position, alongside consonantal modifications to the coronal in verbs such as *mentir* (with modern 1sg *eu minto*), which demonstrate occasional alternative 1sg forms such as *menso*. The existence of diphthongization in some forms and consonantal modification in others, all within 1sg forms of the third conjugation, suggests that in some sense these two processes are 'competing' and that the presence of a coronal consonant can block or impede vowel harmony triggered by the theme vowel. It remains an open question why such a restriction should be generalized, in the case of defective verbs, to other coronals (in fact, only the sonorants), and we speculate that this may have to do with the nature of the theme vowel itself, as a [coronal] vowel under the model of Clements & Hume (1995), and indeed a sonorant as well. The robustness of a phonologically-based generalized points to a way out of the poverty-of-the-stimulus prolem, as the forms that would be ungrammatical are well-circumscribed in terms of generalizable properties -- the very essence of learning. Further work is needed, however, to understand why the 2nd conjugation theme vowel *–e*, also presumably [coronal], would not incur defectiveness in the same way.

Interestingly, our findings demonstrate that the verbs shown in green *emergir, engolir, exaurir*, *explodir, florir, grunhir, latir* and *ungir* are not defective, i.e. not significantly different in their inflected 1pl ind vs. 3sg.sbj forms. Now that we have found *which* the set of defective verbs in Portuguese actually are, in future work, we can examine whether the phonologically-based generalization is extended to novel verbs or not.

In terms of a characterization of defective verbs, we can implement them in terms of the Distributed Morphology operation of Vocabulary Insertion, and state that verbs such as *abolir* have an incomplete list of vocabulary entries. More specifically, suppose that verbs, as open-class lexical items, are represented in the syntax by a root lacking phonological content, call it √abl43 (a notion adopted in Harley 2014). In Arregi & Nevins (2014), it was proposed that the arrangement of vocabulary entries would be as follows for defective verbs such as *abolir* in Spanish and Portuguese:

(5)
√abl43 ↔ /abol/ in the environment: __ i
√abl43 ↔ no elsewhere item

This implementation is in fact similar to the one proposed in Stump (2010:204), where defectiveness is implemented in terms of restrictions on the content to form mapping: a certain set of forms allow such a mapping, and others are simply blocked from it. In a similar manner, for the Surmiran Rumantsch case discussed in Anderson (2010), the verb *dueir* 'should' (from Latin *de+habe:re*) has only non-rhizotonic allomorphs, and thus lacks the 1sg, 2sg, 3sg, 3pl, and present subjunctive forms. (Anderson notes that the defective forms are sometimes substituted by the semantically similar modal *stueir*.). The intuition he provides is that, similarly to the Iberian cases, a morphologization of stress-conditioned vowel alternations has led to a set of rhizotonic and non-rhizotonic allomorphs. In a sense, phonologically-conditioned suppletive allomorphy can give way to defectiveness when there *is* no allomorph outside of the circumscribed environments.

While having a coronal sonorant as the stem-final vowel is not a sufficient condition for defectivity, it is a necessary one, and perhaps such knowledge exactly what is incorporated into the generalization of schemas like (5) to verbs that lack allomorphs outside of this very specific context.

4. CONCLUSION AND FUTURE STEPS

These studies have largely underscored the importance of morphology not-by-itself, but rather, paralleling 'grounded phonology', firmly informed by properties of the interfaces with syntactic structure, phonological patterning, and internal patterns of asymmetrically-organized featural decomposition.

In this work, we have attempted to demonstrate the actual list of defective verbs for real Portuguese speakers, thereby arriving at a closer delimitation than what is noted in dictionaries and descriptive grammars. We have also shown that most verbs which show defectivity according to our task and results have indeed a coronal sonorant following the stem-final vowel, in accordance with Postma's generalization. Thus, the next step is to conduct a wug-test experiment, in order to verify whether the coronal sonorant generalization is productive among verbs that participants have truly never heard.

REFERENCES

ALBRIGHT, Adam. 2003. A quantitative study of Spanish paradigm gaps. In G. Garding and M. Tsujimura (Eds.), *WCCFL 22 Proceedings*, Somerville, MA, pp. 1–14. Cascadilla Press.

ALBRIGHT, Adam. 2009. Lexical and morphological conditioning of paradigm gaps. In *Modeling ungrammaticality*, ed. Curt Rice and Sylvia Blaho, 117–164. London: Equinox Publishing.

ANDERSON, Stephen. 1992. *A-morphous Morphology*. Cambridge University Press.

ANDERSON, Stephen. 2010. Failing one's obligations: defectiveness in Rumantsch reflexes. In. In M. Baerman (ed.) *Defective Paradigms*. London: British Academy/OUP. pp.19-34.

ARREGI, Karlos & Nevins, Andrew. (2014). A monoradical approach to some cases of disuppletion. *Theoretical Linguistics* 40(3/4): 311-330.

ARONOFF, Mark. 1994. *Morphology by Itself.* Cambridge, MA: MIT Press.

CINTRA, Lindley; Cunha, Celso. 2013. *Nova Gramática do Português Contemporâneo* - 6ª Ed., Rio de Janeiro: Lexicon Editorial.

CLEMENTS, G.N. & Elizabeth Hume. 1995. The Internal Organization of Speech Sounds. In John Goldsmith (ed.) *The Handbook of Phonological Theory*, Blackwell, pp. 245-206.

DAMULAKIS, Gean & Marcella Rodrigues. 2013. Alternativas a harmonia vocalica em português arcaico. Paper presented at SIS-Vogais IV, Recife.

HARLEY, Heidi. (2014). On the identity of roots. *Theoretical Linguistics* 40(3/4): 225-276.

MAIDEN, Martin & Paul O'Neill. 2010. 'On morphomic defectiveness: evidence from the Romance languages of the Iberian peninsula'. In M. Baerman (ed.) *Defective Paradigms*. London: British Academy/OUP, pp. 103-124.

POSTMA, Gertjan. 2013. Metaphonic Blocking in Portuguese as a Linearization Deadlock. Workshop on Metaphony, Meertens Institute.

STUMP, Gregory. 2010. Interactions between defectiveness and syncretism. In M. Baerman (ed.) *Defective Paradigms*. London: British Academy/OUP. pp. 181-210.